基于弱监督下改进的CBAM-ResNet18模型识别苹果多种叶部病害

张文景[1,2]，蒋泽中[1]，秦立峰[1,3*]

（1. 西北农林科技大学 机械与电子工程学院，陕西杨凌 712100；2. 农业农村部农业物联网重点实验室，陕西杨凌 712100；3. 陕西省农业信息感知与智能服务重点实验室，陕西杨凌 712100）

摘　要：针对苹果叶部病害图像在仅有图像类别标注的弱监督的条件下识别准确率低的问题，提出了一种基于改进的CBAM-ResNet算法进行苹果叶部病害识别。以ResNet18作为基础模型，对轻量级卷积块注意力模块（Convolutional Block Attention Module，CBAM）注意力机制中通道注意力模块中的多层感知机（Multilayer Perceptron，MLP）进行升维改进，放大苹果叶部病害特征细节；将改进的CBAM融入残差模块中，以加强对关键细节特征的提取，将AlphaDropout配合SeLU（Scaled Exponential Linearunits）融入网络中，防止其网络的过拟合化，加速模型收敛效果；最后，采用单周期余弦退火算法调整学习率，得到病害识别模型。训练在样本图像均只进行图像级标注的弱监督下进行，大大降低标注成本。通过消融实验，探究出改进CBAM中MLP最佳升维维度为2，相对于原CBAM，准确率提升0.32%，并在参数量增加17.59%的情况下，每轮训练时长减少8 s。在包含苹果斑点落叶病、褐斑病、花叶病、灰斑病、锈病等5种病害的6185幅图像数据集上进行了试验测试，结果显示，在弱监督学习下，识别准确率方面，该模型对苹果5种病害的平均识别准确率达到98.44%，改进的CBAM-ResNet18相比改进前的ResNet18提高了1.47%，且高于VGG16，DesNet121，ResNet50，ResNeXt50，EfficientNet-B0和Xception对照模型；在学习效率方面，改进的CBAM-ResNet18相对于ResNet18在参数量增加24.9%的条件下，每轮的训练时间减少6 s，且在VGG16，DesNet121，ResNet50，ResNeXt50，EfficientNet-B0和Xception对照模型中以每轮137 s最快速度完成模型训练。通过混淆矩阵结果，计算出模型的精确度平均值、召回率平均值和F_1分数平均值分别达到了98.43%、98.46%和0.9845。该结果表明，改进的CBAM-ResNet模型可进行苹果叶部病害识别且具有良好的识别结果，可为苹果叶部病害智能识别提供技术支撑。

关键词：病害识别；残差网络；注意力机制；余弦退火学习率；迁移学习；卷积块注意力模块；多层感知机

中图分类号：TP18;S432　　**文献标志码：**A　　**文章编号：**SA202301005

引用格式：张文景，蒋泽中，秦立峰. 基于弱监督下改进的CBAM-ResNet18模型识别苹果多种叶部病害[J]. 智慧农业(中英文)，2023, 5(1): 111-121.

ZHANG Wenjing, JIANG Zezhong, QIN Lifeng. Identifying multiple apple leaf diseases based on the improved CBAM-ResNet18 model under weak supervision[J]. Smart Agriculture, 2023, 5(1): 111-121.

1 引　言

病害是造成苹果减产、降质、降低商品价值的重要因素，及时准确识别病害对及时防治、减少经济损失具有重要意义。早期的病害识别主要是通过人工现场检查完成，费时费力，易受主观影响，因而识别准确率和效率较低。基于数字图像处理和机器学习的病害识别技术极大地推动了病害检测、识别与诊断水平。但传统机器学习技术中由于图像特征提取困难，在面对复杂背景时，算法鲁棒性不

收稿日期：2023-01-09

基金项目：陕西省科学技术研究发展计划项目（2020NY-101）

作者简介：张文景，本科，研究方向为模式识别。E-mail：1418454277@qq.com

*通信作者：秦立峰，博士，副教授，研究方向为农业信息化技术、图像处理与模式识别等。E-mail：fuser@nwafu.edu.cn

足，精度不高。

近年来，深度学习技术在作物病害识别中得到广泛研究并取得了很好的效果。Zhu等[1]利用Inception V2的批量归一化（Batch Normalization，BN）为区域提议网络（Region Proposal Network，RPN）提供多尺度图像特征，在复杂的背景环境下进行植物叶子识别，取得良好效果。丁永军等[2]以VGG-16模型为基础构建卷积胶囊网络，提高了百合病害诊断模型的抗噪能力。Li等[3]为构建植物病虫害视频检测系统，提出了一种基于深度学习的自定义主干突出信息区域以提高模型的识别能力，模型具有良好的鲁棒性。周巧黎等[4]提出了基于改进MobileNetV3的番茄叶片病害分类识别方法，实现番茄病害的实时无损伤检测。深度学习虽然在识别准确率上取得了明显优势，但需以大量精确标注的图像为基础。而在病害识别任务中，由于受到专业领域知识、复杂背景和目标多样等因素影响，标注成本非常高[5]；而无监督学习往往难以在复杂背景下取得良好结果[6]。如何利用实际中存在大量不精确标注的样本，进行模型学习并获得足够高的识别性能是当前亟待解决的问题[7]。在此背景下，弱监督学习[8,9]受到较多关注。在图像识别任务中，弱监督学习只要求标注图像层次的信息，即只标注图像类别而不标注图像中目标的具体位置，即可训练模型并进行识别。Durand等[10]提出的深度卷积网络的弱监督学习方法（Weakly supervised Learning of Deep Convolutional Neural Networks，WILDCAT），在空间性不变约束下，学习多层次局部强化特征，在图像分类、弱监督对象定位和语义分割等3个视觉识别任务上取得良好效果。

弱监督学习的关键问题是在缺少像素级的标注下如何有效关注到图像中的特征。由于注意力机制可自动聚焦图像中重要区域的学习，通过为各通道分配不同权重，使模型学习更加灵活，因此被广泛用于弱监督学习中。Choe和Shim[11]提出了一个基于注意力的（Attention-Based Dropout Layer，ADL），利用自注意力机制处理模型特征图，突出信息区域以提高模型的识别能力。在判定定位的深度特征方面，Zhou等[12]提出对全局平均池化层进行调整，使得卷积神经网络保留其优异的定位性能。对于苹果病害识别问题，虽然强监督学习网络对病害的识别获得较好的效果[13]，但标注成本高，病害样本少。现有的弱监督学习对复杂背景的苹果病害图像仍无法快速有效学习到叶部病害细节特征，存在学习效率低、不易收敛的问题。

针对像素级高精度标注数据集缺乏，而目前弱监督下无法有效关注密集的苹果叶部病害小目标，导致模型学习效率低、收敛不稳定的问题，本文提出了改进的CBAM-ResNet18模型，该模型采用升维多层感知机（Multilayer Perceptron，MLP）改进卷积块注意力模块（Convolutional Block Attention Module，CBAM）中的通道注意力机制，并将改进的CBAM引入ResNet18模型中，将AlphaDropout模块引入模型同时配合激活函数（Scaled Exponential Linearunits，SeLU）使用防止模型过拟合，采用余弦退火学习率调整学习率，在仅利用图像级的类别标签的弱监督下训练模型，实现了对苹果叶部病害的准确识别。

2 试验材料与方法

2.1 苹果叶部病害图像获取

苹果叶部的斑点落叶病、褐斑病、花叶病、灰斑病、锈病的出现数量较多，分布范围较广，在苹果叶部的分布尺寸小且十分密集，对细节特征提取的要求较高，识别难度较高，故本研究选用这5种苹果叶部病害作为识别对象。研究所用图像数据一部分来自2019年周敏敏[14]在西北农林科技大学白水苹果试验站（109°33′32″N，35°12′45″E）、洛川苹果试验站（109°22′35″N，35°47′25″E）和庆城苹果试验站（107°55′36″N，36°0′30″E）3个站点采集的苹果叶面病害图像，分别在晴天、阴天、雨天等不同条件下共采集3217幅图像，分辨率像素为512×512；另一部分于2022年7～10月在陕西省乾县某苹果园区（108°6′33″N，34°27′28″E）利用iPhone11手机在自然光照下采集像素为512×512的图像共1237幅。两个数据集均包含单一和复杂背景下的苹果斑点落叶病、褐斑病、花叶病、灰斑病、锈病等5种常见病害，如图1所示。



(a)斑点落叶病 (b)褐斑病 (c)花叶病 (d)灰斑病 (e)锈病

图1 5种苹果叶部病害图像

Fig. 1 Images of five apple leaf diseases

2.2 病害图像预处理

2.2.1 数据增强

为提高模型的泛化能力、在复杂背景下的抗干扰能力，以及避免训练过拟合，对图像样本进行数据增强[15]。在周敏敏[14]采集的预训练数据集中，采用了两次图像旋转，一次水平和垂直镜像，两次锐度值、两次亮度值和两次对比度调整，一次高斯模糊一共11种数据增强方式，最终预训练集中共有24,348幅病害图片。对在陕西省乾县某苹果园拍摄的图片采用转动、偏移、翻转和亮度变化4类数据增强方式对图像进行扩充，得到6185幅病害照片。最终得到一个共计30,533幅图像的数据集，统一裁剪为224×224像素大小的图像，用于模型训练和测试。病害图像划分的具体情况如表1所示，由于在周敏敏[14]采集的数据集上训练所得模型在实地拍摄中无法获得良好的效果，为对其进行改进，本研究采用迁移学习将周敏敏[14]采集的数据集作为预训练样本，并将其按4∶1划分为预训练集与验证集。陕西省乾县某苹果园拍摄的图片作为测试样本，将其按照3∶1划分为迁移训练集与测试集。

表1 苹果叶部病害数据

Table 1 Apple leaf disease data

病害类别	预训练病害数据集				训练病害数据集			
	原始总数	增强总数	训练集	验证集	原始总数	增强总数	训练集	测试集
斑点落叶病	411	4932	3700	1232	248	1240	927	313
褐斑病	435	5220	3915	1305	258	1290	977	313
花叶病	375	4500	3375	1125	229	1145	861	284
灰斑病	370	4440	3330	1110	250	1250	936	314
锈病	438	5256	3942	1314	252	1260	943	317
总数	2029	24,348	18,262	6086	1237	6185	4644	1541

2.2.2 图像标注

本研究对图像只进行图像类别标注，不对图像进行像素级标注，将每种病害图像归为同一类型样本，从而降低标注成本。

3 病害识别模型构建

3.1 ResNet模型

针对网络易出现梯度消失和梯度爆炸的问题，2015年何恺明等[16]提出了在跳跃连接和恒等映射基础上的残差网络（Residual Network，ResNet），如图2所示。

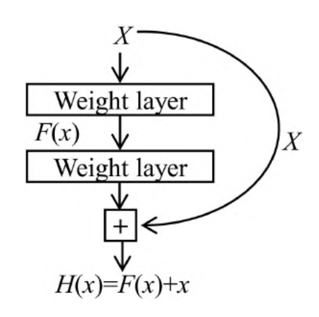


图2 ResNet残差结构

Fig. 2 ResNet residual structure

x为输入信号，经过二次线性变化得到H（x）。在模型学习过程中，输出信号H（x）趋于稳定，且原本带权值的x变为恒等映射，使得输出信号H（x）与输入信号x相等[17]。

对于苹果叶部病害的识别，考虑到大多数场景需要在移动设备上短时间正确识别病害，应采用参数量较小的残差网络模型，故本研究采用ResNet18。

3.2 CBAM注意力机制及其改进

传统卷积神经网络提取的特征不加区分地传输

至下一层，因而无法注意到所需有效信息。尤其在弱监督的条件下。针对此问题，本文在卷积神经网络中融入注意力机制。

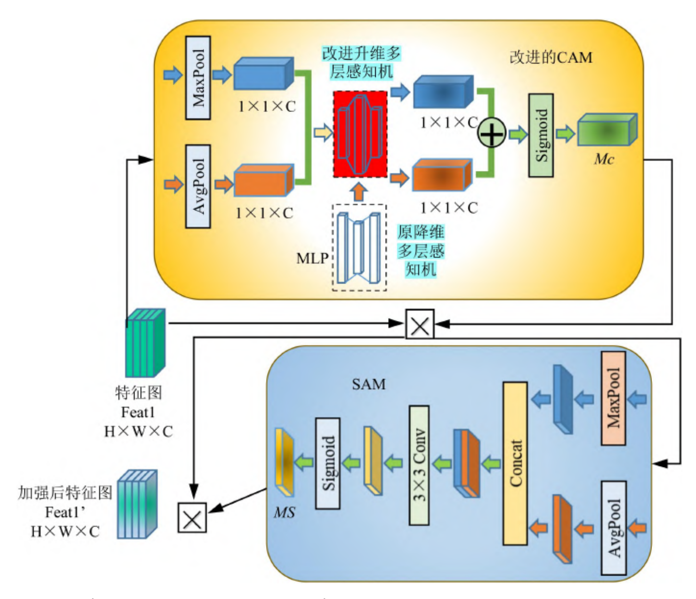
CBAM[18]由通道注意力机制（Channel Attention Module，CAM）[19]结合空间注意力机制（Spartial Attention Module，SAM）[20]组成。它首先对特征图U进行平均池化和全局池化，使其压缩为1×1×C格式，接着将两个压缩特征图送入含有两层神经网络的MLP中，在这两层的神经网络的中间含有一个隐藏层，使压缩后的1×1×C特征图降维至1×1×C/r，将经过MLP的输出进行基于element-wise的加和操作，最终通过Sigmoid函数产生通道注意力特征权重，使其作用于特征图U中产生新的特征图。空间注意力机制能够实现图像空间中关键特征信息的提取，是通道注意力的补充。

由于苹果叶部病害尺寸小且分布十分密集，准确快速的识别对细节特征提取要求较高。为了让CBAM更好注意到图像病害的细节，本研究将CBAM通道注意力模块中的共享神经网络进行改进，如图3所示。原共享神经网络是先将1×1×C的图片降维至1×1×C/r后再升维至1×1×C的特征图，但由于降维对细节损失较大，对最终生成的通道注意力特征权重产生影响，无法重视细节特征。改进的CBAM注意力机制改为先升维再降维，即先将1×1×C的图片升维至1×1×rC后再降维至1×1×C的特征图，提升CBAM注意力机制对苹果叶部病害的分辨能力。

3.3 SeLU激活函数

当模型过大、参数过多时，模型训练容易过拟合。针对此问题，引入Dropout使模型达到正则化的效果。Dropout是让一部分神经元失活以提高模型的泛化能力，但每经过一次Dropout，激活值的分布可能发生变化。对此，Klambauer等[21]提出了AlphaDropout，和一种新的激活函数SeLU，能保持输入与输出的均值和标准差保持一致，归一化性质不变。

相比于激活函数ReLU（Rectified Linear Activation Function）[22]，SeLU不存在失活区，如公式（1），在负无穷存在饱和区，模型的表达不会受



注：Mc表示通道注意力权重;Ms表示空间注意力权重

图3 改进的CBAM注意力机制

Fig. 3 The improved channel CBAM attention mechanism

其影响。经过SeLU激活后，样本分布自动归一化到0均值和单位方差，保证在训练的过程中梯度不会爆炸和消失。

$$SeLU = \lambda \begin{cases} x & , x > 0 \\ \alpha e^{x} - \alpha, & x \leqslant 0 \end{cases} \quad (1)$$

其中，λ和α为超参数，λ约为1.05，αλ约1.67；x为输入量。

本研究将AlphaDropout配合SeLU使用，使特征图输入的分布不发生变化，模型能够更好防止过拟合，提高收敛速度和效果。

3.4 余弦退火算法调整学习率

本研究采用单周期的余弦退火算法[23]衰减学习率，如公式（2）所示，由于本研究为单周期，所以i为1，η_{max}^{i}和η_{min}^{i}分别表示学习率的最大值与最小值，规定了学习率的范围，T_{cur}表示当前执行的轮数，在每轮中的每批训练完进行更新，T_i为总轮数，本研究的预训练阶段T_i为100，训练阶段为50。在开始融入warmup阶段，学习率非常小，防止较大的学习率使模型产生动荡，稳定模型，warmup阶段结束后开始余弦退火优化学习率。

$$\eta_t = \eta_{min}^{i} + \frac{1}{2}(\eta_{max}^{i} - \eta_{min}^{i})(1 + \cos(\frac{T_{cur}}{T_i}\pi)) \quad (2)$$

其中，η_{min}^{i}为学习率最小值；η_{max}^{i}为学习率最大值；T_{cur}为当前执行轮数；T_i为总轮数。

3.5 改进ResNet18的病害识别模型

综合以上，本研究对ResNet18模型进行改进，提出了一种苹果叶部病害识别模型CBAM-ResNet18，结构如图4所示。

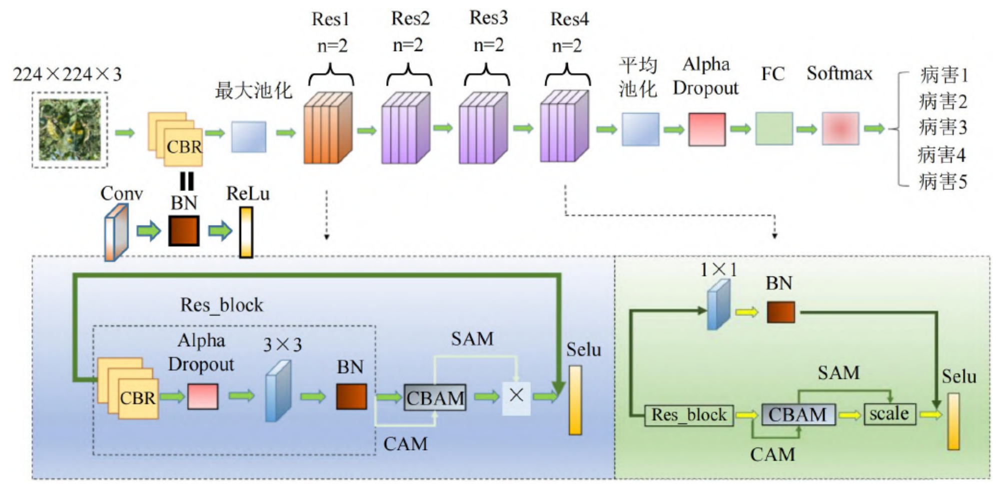


图4 改进的CBAM-ResNet18苹果叶部病害识别模型

Fig. 4 Improved CBAM-ResNet18 apple leaf disease recognition model

在网络的每一个Res模块中融入改进的CBAM注意力机制，以更好地获取病害细节信息，减少复杂背景干扰；将ReLU替换为SeLU并融入AlphaDropout，避免神经元“失活”，防止模型过拟合，保持输入与输出分布一致，使用单周期余弦退火优化算法优化学习率，使模型在收敛过程中保持稳定，避免出现动荡。

改进模型的输入为224×224像素RGB图像。先经卷积层和池化层对图像特征进行初步提取与降维，得到56×56像素的特征图，接着特征图进入4个Res模块进一步提取病害特征[20]。进入Res模块后，首先使用3×3的核对特征图卷积，接着经过激活函数SeLU和AlphaDropout模块，再次进行3×3的卷积，使用归一化模块（Batch Normalization，BN）后，插入CBAM模块产生特征权重；在改进的CBAM通道注意力机制中，特征图先进行通道升维、再降分别降维至1×1×64，1×1×128，1×1×256和1×1×512，获得特征系数，最后与输入特征图相乘得到新的特征图。4个Res模块均使用3×3的卷积核，数量分别为64、128、256和512，特征图的像素尺寸分别为56×56、28×28、14×14和7×7。

在全局平均池化（Global Average Pooling，GloPooling）、全连接（Fully Connected）之中，本研究加入了AlphaDropout模块以防止模型过拟合。使用Softmax函数进行分类，如公式（3）所示，将概率最大值作为结果输出，完成苹果叶部病害的识别。

$$\text{Softmax}(x_i)=\frac{e^{x_i}}{\sum_i^N e^{x_i}}, i\in\{i,\cdots,N\} \tag{3}$$

其中，x_i为输入信号；e^{x_i}为输入信号x_i；$\sum_i^N e^{x_i}$为所有输入信号x_i的指数函数和。

4 试验结果与分析

4.1 模型训练环境与参数设置

模型训练所用CPU配置为Intel®Core™ i7-11800H @2.30 GHz，RTX 3060 Laptop GPU，显存为6 GB，内存为16 GB，基于Windows10操作系统和Tensorflow深度学习框架搭建了深度学习算法训练平台，其中Python版本为3.9.7，Tensorflow版本为2.8.0。

基于迁移学习思想，本研究的模型训练分为预训练和训练两个阶段。预训练阶段在周敏敏[19]数据集上学习并保存最佳模型；之后再利用乾县数据集进行再训练。在预训练阶段，输入图像分辨率为224×224像素；批处理大小（Batch size）为16，所有训练样本迭代一次为一轮，迭代轮数（Epochs）为100轮，采用自适应矩阵估计算法（Adap-

tive Moment Estimation，ADAM）对模型进行优化，损失值采用交叉熵损失函数（Cross Entropy Loss）。为了在保持稳定训练的同时加速模型收敛，初始学习率设置为10^{-6}。经过15个轮迭代的warmup阶段后上升为0.001，使用余弦退火算法调整学习率，最后衰减至0。训练阶段仅将迭代轮数（Epochs）和warmup阶段轮数置为50轮和10轮，其余均与预训练相同。

4.2 消融实验

为了探究在弱监督学习条件下，改进的CBAM共享神经网络升维的最佳维度，同时验证改进CBAM对识别苹果叶部病害的提升效果，AlphaDropout配合激活函数SeLU使用与配合ReLU使用之间的差异，以及单周期余弦退火优化算法与动态指数衰减两种学习率衰减策略对模型学习的影响，进行3组消融实验。

4.2.1 改进的CBAM的影响

为探究改进的CBAM对Resnet18网络的影响，以及通道注意力机制升维的最佳维度，在原Resnet18网络的基础上，进行4组不同MLP的设置，第（1）组为原Resnet18网络中融入原CBAM，第（2）～（4）组分别将CBAM中共享神经网络改进为先升维至1×1×2C、1×1×3C和1×1×5C后，再降维为1×1×C。在相同的数据集上进行训练和测试，结果如表2所示。

表2 CBAM中共享神经网络不同的升维策略下模型的性能

Table 2 The performance of the model under different dimension-up strategies of the shared neural network in CBAM

组号	升维策略	特征图大小	准确率/%	参数量	训练时长/(轮·s^{-1})
（1）	原CBAM	1×1×C/2	98.12	11,889,885	145
（2）	1×1×2	1×1×2C	98.44	13,981,725	137
（3）	1×1×3	1×1×3C	97.14	15,376,285	142
（4）	1×1×5	1×1×5C	97.99	18,165,405	140
（5）	不采用CBAM	——	97.99	11,189,893	143

由表2可见，改进的注意力机制将特征图升维至1×1×2C的效果最好，其准确率分别高于其他两种升维方式1.3%和0.45%，较原CBAM注意力机制准确率提升了0.32%；且升维至1×1×2C参数量比对照组（3）低9.97%，比对照组（4）低29.92%，比原CBAM高17.59%情况下，以最快学习速度完成训练，相比于对照组（1）、（3）和（4）分别缩短8 s、5 s和3 s。

以上试验结果表明，将特征图由压缩变为升维，能将图像病害细节放大，从而有效提取图像病害的细节；但若升维过大，也会将复杂背景噪声放大，影响识别。故综合考虑准确率、参数量及学习效率，本研究选用共享神经网络升维1×1×2C的改进注意力机制。

4.2.2 AlphaDropout配合SeLU使用的影响

为验证AlphaDropout配合SeLU使用对ResNet18网络在弱监督学习条件下识别苹果叶部病害的影响，在原ResNet18网络的基础上，分别测试AlphaDropout+ReLU，AlphaDropout +SeLU和原ResNet三种模型，在同样条件下三种模型的识别结果如表3所示，可见与ReLU相比，AlphaDropou+SeLU的准确率最高为97.34%，较AlphaDropou+ReLU提高0.13%，且每轮的训练时长减少2 s。这是因为SeLU有效避免神经元“失活”问题，提高模型表达能力，且其自归一化性质配合AlphaDropout一起使用，保证了输出是0均值和单位标准差，使得模型收敛速度加快。

表3 不同激活函数的模型性能表现

Table 3 Model performances with different activation functions

模型	准确率/%	训练时长/(轮·s^{-1})
AlphaDropou+ReLU	97.21	147
AlphaDropou+SeLU	97.34	145
原模型	96.97	143

4.2.3 单周期的余弦退火优化算法的影响

为探究单周期余弦退火优化算法对本模型的影响，将其与动态指数衰减学习率作对比，观察模型性能表现。

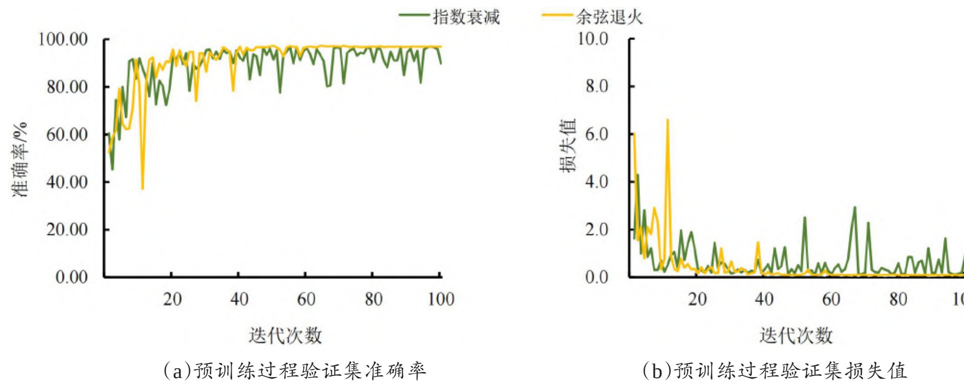
由表4可知，余弦退火优化算法对于学习率的

表4 不同学习率调节方式的模型性能

Table 4 Model performances with different learning rate adjustment methods

模型	准确率/%	训练时长/(轮·s^{-1})
单周期余弦退火	97.99	143
动态指数衰减	97.90	143
固定学习率	96.92	143

调节更符合改进的CBAM-ResNet18模型的学习，在前期warmup阶段能稳定模型，后期按余弦退火优化算法进行调节使模型更好地学习病害的特征，准确率比动态衰减学习率提升0.09%。图5为采用不同的学习率学习验证集的准确率变化曲线。由图5可见，余弦退火算法优化的学习率在后期衰减合理，使得训练过程快速进入收敛且保持较高的准确率。指数衰减学习率在同样的训练轮次后仍有较大波动，模型收敛不好。



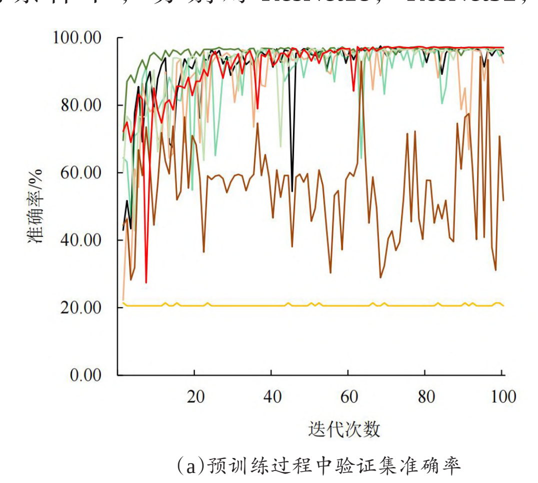
(a)预训练过程验证集准确率　(b)预训练过程验证集损失值

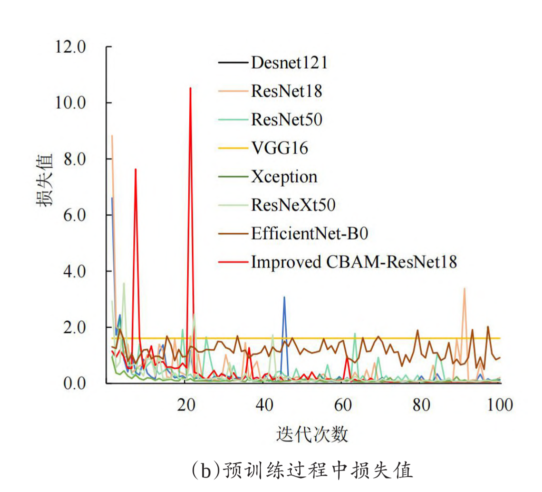
图5 指数衰减和余弦退火算法优化模型的学习过程中验证集准确率变化曲线

Fig. 5 The variation curve of validation set accuracy during the learning process of exponential index decay and cosine annealing algorithm optimization model

4.3 预训练过程对比分析

得出改进模型的最佳参数后，测试改进的CBAM-ResNet18模型预训练过程中性能表现，在同样的条件下，分别对ResNet18，ResNet52，DesNet121，VGG16，Xception，ResNeXt50，EfficientNet-B0和本研究改进的CBAM-ResNet18模型进行预训练，各个模型在预训练后在验证集上的准确率与损失值如图6所示。





(a)预训练过程中验证集准确率　(b)预训练过程中损失值

图6 不同模型在预训练过程中测试集的准确率与损失值变化曲线

Fig. 6 The curve of accuracy and loss value of the test set during the training process of different models

从图6可见，随着训练轮数的增加，除了VGG16模型外的7种模型在验证集上的准确率整体上呈现上升趋势，损失率则下降至一个相对稳定的区间内振荡。而VGG16准确率一直停留21%左右，无法收敛。在收敛后的模型中。EfficientNet-B0波动大，准确率在60%左右且波动起伏过大，收敛效

果不佳。ResNet18和Resnet50在收敛后，模型准确率在96%左右有较大的振荡，ResNet18比Resnet50的振荡更明显，甚至会降到80%~60%，而损失值在0.2左右有较大的振荡。具有多尺度结构的Xception和ResNeXt50收敛后较为稳定，准确率在96%左右伴随较小震荡，损失值稳定在0.1左右。DesNet121准确率和损失值分别稳定在96%和0.1左右，相比于ResNet50较好。这是由于在弱监督的条件下，没有像素级的标注，模型对有效的病害特征学习能力不足，无法稳定其准确率和损失值，产生了较大的波动。相比之下，本研究改进的CBAM-Resnet18模型的曲线相对平滑，波动小于其他模型，表现明显更为稳定。

4.4 病害识别试验及结果分析

为进一步对比模型对于苹果叶部病害识别的效果，将预训练好的模型保存并迁移至训练集图像上进行迁移学习，以调整参数获得目标模型，最后在测试集上进行病害识别实验，测试模型的分类性能。采用准确率、参数量和训练时长对模型性能进行评价，其中准确率如公式（4）所示。试验结果如表5所示。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

其中，TP为被模型预测为正类的正样本个数；TN为被模型预测为负类的负样本个数；FP为被模型预测为正类的负样本个数；FN为被模型预测为负类的正样本个数。

表5 不同模型的苹果叶部病害识别性能表现

Table 5 Performance comparison of different classic networks of apple leaf disease recognition

模型	验证集准确率/%	测试集准确率/%	参数量	训练时长/(轮·s^{-1})
VGG16	21.44	20.51	134,281,029	186
DesNet121	97.12	98.12	7,042,629	188
ResNet50	97.05	98.18	23,597,957	156
ResNeXt50	97.14	97.29	23,084,933	442
ResNet18	96.97	96.97	11,189,893	143
EfficientNet-B0	94.36	90.12	4,055,969	147
Xception	97.26	97.93	20,778,725	234
CBAM-ResNet18	97.38	98.44	13,981,725	137

由表5可见，与VGG16、DesNet121、ResNet50、ResNeXt50、ResNet18、EfficientNet-B0和Xception相比，本研究提出的改进CBAM-ResNet18在测试集上的准确率分别高出77.93%、0.32%、0.26%、1.15%、1.47%、8.32%和0.51%。ResNet50利用残差结构确保其具有很强的特征学习能力，因此可以达到98.18%的准确率。DenseNet121的基本结构Dense Block具有增强特征传播和激励特征重用的优点，使其以ResNet50三分之一的参数获得相近的准确率。此外，ResNeXt50的训练时间达到了442 s/轮，这是由于大量并行分支的存在，大大降低了模型的计算效率。与ResNeXt50类似，Xception也具有多尺度结构，这就导致了模型参数之间的交叉计算，导致计算时间偏高。相比之下，本研究改进的CBAM-ResNet18相对于ResNet18在参数量增加24.9%的情况下，不但未造成以往卷积神经网络参数量的增加带来学习效率降低的负面影响，反而学习效率和速度均得到提升，每轮训练时间缩短6 s。这是由于改进的注意力机制使卷积层能够更迅速地关注到了特征图的病害特征，提高特征学习速率和训练效率，在参数量增加的情况下减少了训练时间。相对于同为残差网络的ResNet50，改进的CBAM-ResNet18保持了残差网络ResNet18的低参数量，每一轮训练时间缩短19 s，且准确率高出0.26%，为部署在更多存储量较小、运算性能配置不高的硬件终端设备提供了可能性。

以上结果表明，对于苹果叶部病害的小且分布密集的特征，改进的CBAM-ResNet18模型在缺少像素级标注的弱监督学习条件下，能够有效地提取

特征，改进的CBAM注意力机制能够强化网络对图像中特定显著区域特征表达能力，带来较其他网络更高的准确率，低参数量，同时大大提高学习病害效率，减少学习时间。

4.5 多类别识别混淆矩阵

混淆矩阵常用来进行多分类性能评估。本研究共涉及5种苹果叶部病害，用改进的CBAM-ResNet18模型对测试集数据进行分类识别实验，得到混淆矩阵如图7所示。

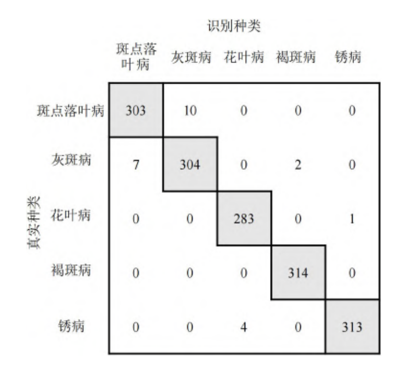


图7 苹果叶部病害测试集的混淆矩阵

Fig. 7 Confusion matrix of apple leaf disease test set

混淆矩阵中判断模型性能的指标通常有精确度（Precision）、召回率（Recall）和 F_1 分数（F_1 score），公式如（5）～（7）。

$$Prescision = \frac{TP}{TP + FP} \times 100\% \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{6}$$

$$F_1 = \frac{2 \times Prescision \times Recall}{Prescision + Recall} \tag{7}$$

其中，TP和FN表示病害样本被正确识别为某种类和非某种类的数量，个；FP和TN表示病害样本被错误识别为某种类和非某种类的数量，个。精确度为所有预测为某类别的样本中预测正确的比例，%；召回率表示所有真实类别中预测正确的比例，%；F_1分数是精确率和召回率的调和平均数，是二者的综合评价。根据图7的混淆矩阵，得到5种苹果叶部病害种类识别试验结果如表6所示。

由表6可知，改进的CBAM-ResNet模型在精确度、召回率和F_1分数上均取得良好的结果。由图7可知，本研究改进的模型最多的误识别为灰斑病与斑点落叶病两种病害，斑点落叶病中有10幅图像被识别错误，且均被误认为成灰斑病，而灰斑病中7幅图像被识别错误，且均被识别为斑点落叶病，这是由于灰斑病在苹果叶片上为黄褐色圆点，斑点落叶病为褐色圆点，两者颜色、形态、大小分布都十分相似，故而易造成混淆。其他3种病害的精确度，召回率和F_1分数均接近100%，识别效果较好；其中有4幅锈病图像被识别为花叶病，1幅花叶图片被识别为锈病，由于密集分布的锈病图像与花叶病图像十分相似，被错误识别；另外，由于褐斑病叶面为绿色叶面，有褐色斑点，因此有2幅灰斑病被错误地识别为小面积的褐斑病。总体误识别数较少，表明改进的CBAM-ResNet模型性能良好，能够适用于实际场景下苹果叶部病害的识别。

表6 5类苹果叶部病害的精确度、召回率、F_1分数的实验结果

Table 6 Experimental results of precision, recall, and F_1 scores for 5 types of apple leaf diseases

病害类	样本数量/幅	精确度/%	召回率/%	F_1分数	错误识别率/%
斑点落叶病	248	97.74	96.81	0.9727	2.26
灰斑病	250	96.82	97.12	0.9697	3.18
花叶病	229	98.61	99.65	0.9913	1.39
褐斑病	258	99.37	100.00	0.9968	0.63
锈病	252	99.60	98.74	0.9921	0.40
平均值	——	98.43	98.46	0.9845	1.57

5 结 论

针对仅有图像标签弱监督学习条件下无法快速有效关注苹果叶部病害图像特征，识别准确率和学习效率低的问题，本研究提出了一种苹果叶部病害识别模型，在ResNet18模型中插入CBAM的基础上，为了更好地关注病害特征，对CBAM中MLP进行改进，先升维以更好地突出了苹果叶部病害的特征细节后，让模型能够更加有效病害特征更好地学习病害特征及细节，再进行降维；同时对激活函数以及学习率进行改进，提高了模型的训练效率，更注重于病害细节的学习。得到以下结论：

（1）以ResNet18为基础模型，在每一个残差块中插入CBAM，采用升维MLP改进CBAM，将先降维后升维的MLP改进为先升维后降维，可有效放大苹果图像病害特征的细节，提升卷积层对于病害特征的提取效率。

（2）在弱监督学习的条件下，改进的CBAM-

ResNet18网络模型对5种苹果叶部病害平均识别准确率达98.44%，不仅在准确率上超过了对照模型（VGG16，DesNet121，ResNet50，ResNeXt50，ResNet18，EfficientNet-B0，Xception），而且相对于原模型ResNet18在参数量增加24.9%的情况下，不但未造成以往卷积神经网络参数量的增加带来学习效率降低，训练时长增加的负面影响，反而学习效率和速度均得到提升，每轮训练时间相对于原模型ResNet18减少了6 s，且高于所有对照模型。

（3）通过混淆矩阵对种类识别结果分析，5种病害的精确度平均值、召回率平均值、F_1分数平均值分别达到了98.43%、98.46%和0.9845，其中花叶病、褐斑病、锈病的精确度平均值、召回率平均值、F_1分数十分接近100%，验证了改进的CBAM-ResNet18模型性能优越，能够有效分辨苹果病害。

（4）虽然基于弱监督学习的条件下能够有效关注到苹果叶部病害的特征细节实现优秀的准确率和学习效率，但本模型的大小与参数量仍有提升空间。同时，考虑建立基于云服务的苹果叶部病害识别诊断服务系统，并开发移动端App，以满足产业应用需求。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] ZHU X L, ZHU M, REN H E. Method of plant leaf recognition based on improved deep convolutional neural network[J]. Cognitive systems research, 2018, 52: 223-233.

[2] 丁永军, 张晶晶, 李民赞. 基于卷积胶囊网络的百合病害识别研究[J]. 农业机械学报, 2020, 51(12): 246-251, 331.
DING Y J, ZHANG J J, LI M Z. Disease detection of lily based on convolutional capsule network[J]. Transactions of the Chinese society for agricultural machinery, 2020, 51(12): 246-251, 331.

[3] LI D S, WANG R J, XIE C J, et al. A recognition method for rice plant diseases and pests video detection based on deep convolutional neural network[J]. Sensors (basel, Switzerland), 2020, 20(3): ID 578.

[4] 周巧黎, 马丽, 曹丽英, 等. 基于改进轻量级卷积神经网络MobileNetV3的番茄叶片病害识别[J]. 智慧农业(中英文), 2022, 4(1): 47-56.
ZHOU Q L, MA L, CAO L Y, et al. Identification of tomato leaf diseases based on improved lightweight convolutional neural networks MobileNetV3[J]. Smart agriculture, 2022, 4(1): 47-56.

[5] 任冬伟, 王旗龙, 魏云超, 等. 视觉弱监督学习研究进展[J]. 中国图象图形学报, 2022, 27(6): 1768-1798.
REN D W, WANG Q L, WEI Y C, et al. Progress in weakly supervised learning for visual understanding[J]. Journal of image and graphics, 2022, 27(6): 1768-1798.

[6] MEI S A, YANG H A, YIN Z P. An unsupervised-learning-based approach for automated defect inspection on textured surfaces[J]. IEEE transactions on instrumentation and measurement, 2018, 67(6): 1266-1277.

[7] 孙美君, 吕超章, 韩亚洪, 等. 弱监督学习下的融合注意力机制的表面缺陷检测[J]. 计算机辅助设计与图形学学报, 2021, 33(6): 920-928.
SUN M J, LYU C Z, HAN Y H, et al. Weakly supervised surface defect detection based on attention mechanism[J]. Journal of computer-aided design & computer graphics, 2021, 33(6): 920-928.

[8] DESELAERS T, ALEXE B, FERRARI V. Weakly supervised localization and learning with generic knowledge[J]. International journal of computer vision, 2012, 100(3): 275-293.

[9] RUSSAKOVSKY O, LIN Y Q, YU K, et al. Object-centric spatial pooling for image classification[C]// European conference on computer vision. Berlin, Heidelberg, Germany: Springer, 2012: 1-15.

[10] DURAND T, MORDAN T, THOME N, et al. WILDCAT: weakly supervised learning of deep ConvNets for image classification, pointwise localization and segmentation[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 5957-5966.

[11] CHOE J, SHIM H. Attention-based dropout layer for weakly supervised object localization[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 2214-2223.

[12] ZHOU B L, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 2921-2929.

[13] 王云露, 吴杰芳, 兰鹏, 等. 基于改进Faster R-CNN的苹果叶部病害识别方法[J]. 林业工程学报, 2022, 7(1): 153-159.
WANG Y L, WU J F, LAN P, et al. Apple disease identification using improved Faster R-CNN[J]. Journal of forestry engineering, 2022, 7(1): 153-159.

[14] 周敏敏. 基于迁移学习的苹果叶面病害Android检测系统研究[D]. 杨凌: 西北农林科技大学, 2019.
ZHOU M M. Apple foliage diseases recognition in android system with transfer learning-based[D]. Yangling: Northwest A & F University, 2019.

[15] 谢秋菊, 吴梦茹, 包军, 等. 融合注意力机制的个体猪脸识别[J]. 农业工程学报, 2022, 38(7): 180-188.
XIE Q J, WU M R, BAO J, et al. Individual pig face recognition combined with attention mechanism[J]. Transactions of the Chinese society of agricultural engineering, 2022, 38(7): 180-188.

[16] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
[17] 涂雪滢, 刘世晶, 钱程. 基于ResNet的典型养殖鱼类识别方法研究[J]. 渔业现代化, 2022, 49(3): 81-88.
TU X Y, LIU S J, QIAN C. Study on the identification methods of typical cultured fish based on ResNet[J]. Fishery modernization, 2022, 49(3): 81-88.
[18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]// Computer Vision-ECCV 2018. Berlin, Heidelberg, Germany: Springer International Publishing, 2018: 3-19.
[19] FU J L, ZHENG H L, MEI T. Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 4476-4484.
[20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.
[21] KLAMBAUER G, UNTERTHINER T, MAYR A, et al. Self-normalizing neural networks[J/OL]. arXiv: 1706.02515, 2017.
[22] GLOROT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks[C]// Artificial Intelligence and Statistics Conference. Cambridge, US: MIT Press, 2011: 315-323.
[23] LOSHCHILOV I, HUTTER F. SGDR: Stochastic gradient descent with warm restarts[J/OL]. arXiv: 1608.03983, 2016.

Identifying Multiple Apple Leaf Diseases Based on the Improved CBAM-ResNet18 Model Under Weak Supervision

ZHANG Wenjing[1,2], JIANG Zezhong[1], QIN Lifeng[1,3*]

(1. *College of Mechanical and Electronic Engineering*, *Northwest A&F University*, *Yangling* 712100, *China*; 2. *Key Laboratory of Agricultural Internet of Things*, *Ministry of Agriculture and Rural Affairs*, *Yangling* 712100, *China*; 3. *Key Laboratory of Agricultural Information Perception and Intelligent Services*, *Yangling* 712100, *China*)

Abstract: To deal with the issues of low accuracy of apple leaf disease images recognition under weak supervision with only image category labeling, an improved CBAM-ResNet-based algorithm was proposed in this research. Using ResNet18 as the base model, the multilayer perceptron (MLP) in the lightweight convolutional block attention module (CBAM) attention mechanism channel was improved by up-dimensioning to amplify the details of apple leaf disease features. The improved CBAM attention module was incorporated into the residual module to enhance the key details of AlphaDropout with SeLU (Scaled Exponential Linearunits) to prevent overfitting of its network and accelerate the convergence effect of the model. Finally, the learning rate was adjusted using a single-cycle cosine annealing algorithm to obtain the disease recognition model. The training test was performed under weak supervision with only image-level annotation of all sample images, which greatly reduced the annotation cost. Through ablation experiments, the best dimensional improvement of MLP in CBAM was explored as 2. Compared with the original CBAM, the accuracy rate was increased by 0.32%, and the training time of each round was reduced by 8 s when the number of parameters increased by 17.59%. Tests were conducted on a dataset of 6185 images containing five diseases, including apple spotted leaf drop, brown spot, mosaic, gray spot, and rust, and the results showed that the model achieved an average recognition accuracy of 98.44% for the five apple diseases under weakly supervised learning. The improved CBAM-ResNet18 had increased by 1.47% compared with the pre-improved ResNet18, and was higher than VGG16, DesNet121, ResNet50, ResNeXt50, EfficientNet-B0 and Xception control model. In terms of learning efficiency, the improved CBAM-ResNet18 compared to ResNet18 reduced the training time of each round by 6 s under the condition that the number of parameters increased by 24.9%, and completed model training at the fastest speed of 137 s per round in VGG16, DesNet121, ResNet50, ResNeXt50, Efficient Net-B0 and Xception control models. Through the results of the confusion matrix, the average precision, average recall rate, and average F_1 score of the model were calculated to reach 98.43%, 98.46%, and 0.9845, respectively. The results showed that the proposed improved CBAM-ResNet18 model could perform apple leaf disease identification and had good identification results, and could provide technical support for intelligent apple leaf disease identification providing.

Key words: disease identification; residual network; attentional mechanisms; cosine annealing learning rate; transfer learning; convolutional block attention module (CBAM); multilayer perceptron (MLP)